



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

WALTTERI VUORIMIES
URHEILUANALYTIIKAN HYÖDYNTÄMINEN KORIPALLOSSA

Kandidaatintyö

Tarkastaja: Pasi Hellsten

TIIVISTELMÄ

Waltteri Vuorimies: Urheiluanalytiikan hyödyntäminen koripallossa

Utilizing sport analytics in basketball

Tampereen teknillinen yliopisto

Kandidaatintyö, 21 sivua

Maaliskuu 2018

Tietojohtamisen kandidaatin tutkinto-ohjelma

Pääaine: Tietojohtaminen

Tarkastaja: Pasi Hellsten

Avainsanat: kandidaatintyö, urheiluanalytiikka, koripallo, hyödyntäminen

Tutkimusongelmana työssä on eksplisiittisen tiedon hyödyntämisen vähäisyys koripallo urheilusuorituksen parantamisessa, vaikka edellytykset ovat olleet kunnossa pidemmän aikaa. Aihe kuuluu analytiikkaan, keskittyen juuri urheiluanalytiikkaan koripallossa. Analytiikka on osa tietojohtamisen aluetta, joten aihe sopii hyvin kandidaatintyön kohteeksi. Tutkimusongelmaan liittyy vahvasti se, että keinot hankkia dataa koripallossa ovat kunnossa ja dataa on paljon saatavilla joukkueilla. Datan määrään nähden, siitä jalostettua eksplisiittistä tietoa ei kuitenkaan hyödynnetä tarpeeksi. Tutkimus käsittelee pääosin NBA-joukkueiden urheiluanalytiikan menetelmiä sekä mahdollisuuksia, koska NBA-sarja on maailman suurin ja kova tasoisin koripallosarja. Näin ollen tämän sarjan joukkueilla on resursseja urheiluanalytiikan hyödyntämiseen enemmän kuin esimerkiksi suomalaisilla joukkueilla.

Tutkimuksen tuloksena löydettiin useita eri keinoja hyödyntää eksplisiittistä tietoa koripalloilussa. Dataa saadaan useista lähteistä ja tästä jalostetun eksplisiittisen tiedon hyödyntämisen keinoja on lukuisia. Merkittävimpiä tiedon lähteitä urheiluanalytiikan osalta koripallossa ovat ottelutilastot sekä Sport VU -kameroiden avulla hankittu data. Modernimpia keinoja hankkia dataa ovat työn tuloksien perusteella sosiaalisen median tutkiminen sekä puettavan teknologian hyödyntäminen. Merkittävänä tuloksena voidaan työssä pitää myös sitä, että eri tutkimuksissa on huomattu perinteisten pelaajien roolien olevan virheellinen tapa pelaajien vertailuun.

ALKUSANAT

Urheilu on aina ollut intohimoni ja yhdistämällä koulutukseni sekä kiinnostuneisuuteni urheiluun syntyi ajatus urheiluanalytiikan tutkimisesta. Urheiluanalytiikka on myös ajankohtainen aihe, mikä lisäsi mielenkiintoani tutkimusta kohtaan. Urheiluanalytiikkaan perehtyessä opin näkemään urheilun eri näkökulmasta ja sain paljon uutta tietoa aiheesta.

Kandidaatintyö tehtiin Tampereen Teknisessä Yliopistossa, tietojohdamisen koulutusohjelmaan. Työn tekemiseen sain apua ja neuvoja Pasi Hellsteniltä sekä kandidaatintyökurssin ryhmäläisiltäni. Haluan kiittää heitä rakentavasta palautteesta sekä auttamisesta muun muassa aiheen rajauksessa.

Tampereella, 18.4.2018

Waltteri Vuorimies

SISÄLLYSLUETTELO

1.	JOHDANTO.....	1
1.1	Tutkimuksen rajausta ja tavoitteet	2
1.2	Tutkimusongelma ja työn rakenne	2
2.	TUTKIMUSMENETELMÄ JA -AINEISTO	4
2.1	Tutkimusmenetelmä	4
2.2	Tutkimusaineisto	4
3.	URHEILUANALYTIikka.....	6
3.1	Datan hankinta ja jalostus	6
3.2	Tiedon hyödyntäminen.....	8
3.3	Urheiluanalytiikan haasteet	9
4.	EKSPLISIITTISEN TIEDON HYÖDYNTÄMINEN KORIPALLOSSA	12
4.1	Ottelutilastot	12
4.2	Seurantadata	13
4.3	Roolien muuttuminen.....	14
4.4	Teknologian tuomat mahdollisuudet	15
4.5	Sosiaalinen media.....	15
5.	YHTEENVETO	17
5.1	Yhteenveto	17
5.2	Tulosten arviointi ja jatkotutkimusmahdollisuudet.....	18
	LÄHTEET	19

1. JOHDANTO

Urheiluanalytiikka on tällä hetkellä suurena trendinä, vaikka sitä on hyödynnetty koripallonkin osalta jo vuosia (Sellitto 2015). Pienessä määrin urheiluanalytiikka on ollut olemassa koko urheilun historian ajan. Jo vuonna 1860-luvulla on kerätty yksinkertaisia tilastoja ottelutapahtumista (Gudmundsson & Horton 2017). Nykyaikana modernit teknologiat mahdollistavat uudenlaisia datan keräystapoja, sekä uudenlaisia keinoja hyödyntää datasta jalostettua eksplisiittistä tietoa. Näitä uudenlaisia menetelmiä kerätä dataa ovat muun muassa vuonna 2014 kaikkien NBA-joukkueiden halleihin tulleet Sport VU -kamerat, jotka keräävät 25 datapistettä sekunnissa pelaajista ja pallosta (Sellitto 2015).

Urheiluanalytiikka sisältää datan keräyksen, sen jalostamisen hyödynnettävään muotoon ja jalostetun datan hyödyntämisen eli eksplisiittisen tiedon hyödyntämisen. Urheiluanalytiikkaa hyödynnetään, jotta saadaan etua kilpailijoihin nähden. Urheiluanalytiikan avulla valmennus ja muu johto saavat dataa päätöksenteon tueksi. (Shields 2016) Tämä tutkimus keskittyy koripalloilussa käytettävään urheiluanalytiikkaan, pääosin NBA-sarjassa hyödynnettyyn urheiluanalytiikkaan. Tutkimus on mielenkiintoista tehdä juuri koripallosta, sillä se on baseballin ohella yksi urheiluanalytiikkaa eniten hyödyntävistä urheilulajeista (Martinez et al. 2017). NBA-sarja on tutkimuksen keskiössä siitä syystä, että se on maailman suurin ja kovatasoisin koripallosarja, ja tämän takia urheiluanalytiikkaa käytetään siinä sarjassa eniten.

Koripalloilun osalta urheiluanalytiikka on pitkään ollut pelkkien ottelutilastojen kirjaimista ja niiden pohjalta tehtyjen johtopäätösten analysointia. Näitä ottelutilastoja kerätään vieläkin ja niitä hyödynnetään, mutta niiden rinnalle on tullut muitakin analytiikan välineitä. (van Bommel & Bornn 2017) Ottelutilastojen pohjalta tehtyjen analyysien yksi vaikeus on muun muassa se, että ne eivät kerro pelaajien todellista arvoa (Deshpande & Jensen 2016). Perinteiset tilastot kertovat esimerkiksi sen, kuka on tehnyt pisteitä ja kuka on ottanut levypalloja, mutta eivät kerro sitä, milloin ja missä tilanteessa kyseiset tilastomerkinnot ovat tehty (Deshpande & Jensen 2016).

Urheiluanalytiikan teoriasta ja sen hyödyntämisestä kerrotaan myöhemmässä vaiheessa tarkemmin. Tämän johdannon tarkoituksena oli kuitenkin luoda lukijalle jonkinlainen mielikuva siitä, mitä urheiluanalytiikka on ja etenkin siitä, mitä se koripalloilun osalta kattaa. Lopuksi todetaan vielä, että urheiluanalytiikka voidaan soveltaa moneen hyvin erilaiseen tarkoitukseen koripallon osalta. Urheiluanalytiikka on kuitenkin vielä kehitysvaiheessa, joten myös haasteita urheiluanalytiikan osalta löytyy. Seuraavissa ala-

luvuissa käsitellään siis tutkimuksen rajaus ja tavoitteet sekä tutkimusongelma ja työn rakenne.

1.1 Tutkimuksen rajaus ja tavoitteet

Tutkimuksen aihe kuuluu liiketoimintatiedon hallinnan osa-alueeseen, joka on rajattu käsittelemään liiketoimintatiedon hallintaa koripallossa, toisin sanoen urheiluanalytiikka koripallossa. Aihe kattaa sen, miten ja mitä eksplisiittistä tietoa koripallossa hankitaan, miten sitä jalostetaan sekä miten sitä hyödynnetään. Tutkimus käsittelee koripalloilun osalta ainoastaan suorituksen parantamisen näkökulmaa, joten esimerkiksi vedonlyönnin sekä liiketoiminnan näkökulmat ovat jätetty pois tutkimuksesta. Tämä kandidaatin työ painottuu eniten eksplisiittisen tiedon hyödyntämiseen koripallossa, mutta datan hankintaa ja jalostusta tulee myös käsitellä, jotta tutkimus muodostaa kattavan kokonaiskuvan. Hiljaisen tiedon hyödyntäminen on myös rajattu aiheesta pois, koska tästä osa-alueesta ei ole kirjallisuutta eikä sen sisällyttäminen kandidaatintyöhön olisi laajuuden puitteissa mahdollista.

Tutkimus keskittyy analytiikan osalta urheilusuurituksen parantamiseen koripallon pelaamisessa, koska tämä rajaus itsessään on jo tarpeeksi laaja kandidaatintyön alueeksi eli muiden näkökulmien mukaan ottaessa työn koko kasvaisi liian suureksi. Tutkimus on rajattu tähän spesifiin osa-alueeseen myös sen takia, että kilpaurheilun päätavoitteena on voittaa vastustaja eli näin ollen yksilön tai joukkueen suorituksen parantaminen on tärkeää. Tutkimuksen rajauksen kannalta on tärkeää myös huomata, että kyseessä on ainoastaan yksiläji, joka on koripallo, eikä eksplisiittisen tiedon hyödyntämistä tutkita tässä työssä yleisesti kaikkien lajien osalta.

Tutkimuksen tavoitteena on löytää vastauksia tutkimuskysymyksiin sekä luoda kattava kuva etenkin siitä, millä tavalla ja mihin eksplisiittistä tietoa hyödynnetään koripallosuorituksen parantamisessa. Pää- ja alatutkimuskysymykset sekä tutkimusongelma esitellään seuraavassa alaluvussa. Tavoitteena työssä on myös tutkia urheiluanalytiikan hyödyntämisen kehitystä vuosien varrelta sekä katsastaa, mihin eksplisiittistä tietoa voitaisi nykyteknologian puitteissa hyödyntää.

1.2 Tutkimusongelma ja työn rakenne

Tutkimusongelmana työ käsittelee eksplisiittisen tiedon hyödyntämisen puutetta, vaikka tarvittava data on saatavilla. Eksplisiittistä tietoa on hyödynnetty koripallossa monia vuosia vähäisellä tasolla (Gudmundsson & Horton 2017), mutta suuri murros on vasta nyt käynnissä ja uudet tekniikat mahdollistavat eksplisiittisen tiedon hyödyntämisen monella tapaa. Uusien teknologioiden sekä heränneen kiinnostuksen vuoksi eksplisiittis-

tä tietoa tulisi hyödyntää, ja sitä hyödynnetään koripallossa koko ajan enemmän (Vinué & Epifanio 2017).

Tutkimusongelmaa ratkaistaan tässä kandidaatintyössä tutkimuskysymysten avulla. Pää-tutkimuskysymys työssä on:

- Miten urheiluanalytiikkaa voidaan hyödyntää urheilusuorituksen parantamisessa koripallossa?

Päättutkimuskysymykseen vastataan alatutkimuskysymyksillä, jotka ovat:

- Miten koripalloseurat hankkivat tarvittavaa dataa?
- Mihin koripalloseurat hyödyntävät eksplisiittistä tietoa urheilusuorituksen parantamisen lisäksi?
- Mitä haasteita urheiluanalytiikkaan liittyy koripallon pelaamisen osalta?

Päättutkimuskysymys kattaa laajasti kaikki keinot ja tavat, joiden avulla urheiluanalytiikkaa hyödynnetään koripallossa. Päättutkimuskysymyksen avulla ei ainoastaan tutkita sitä, mihin tällä hetkellä eksplisiittistä tietoa hyödynnetään, vaan se kattaa myös urheiluanalytiikan kohteet, joissa eksplisiittistä tietoa ei vielä hyödynnetä. Alatutkimuskysymyksien avulla päättutkimuskysymys on jaettu pienempiin osiin, joka helpottaa aiheen tutkimista eri näkökulmista. Työn aiheen ollessa eksplisiittisen tiedon hyödyntäminen, on tärkeää tutkia myös, millaista tietoa hyödynnetään ja mihin sekä miksi tietoa ylipäätään hyödynnetään koripallossa.

Työn rakenne muodostuu niin että, tämän johdanto luvun jälkeen, luvussa 2 käydään läpi tutkimusmenetelmä ja –aineisto eli esitellään, miten tutkimus on toteutettu. Luvussa 3 käydään läpi yleisesti mitä urheiluanalytiikka on ja mitä ongelmia sen hyödyntämisessä on. Luvussa 3 esitellään myös kaikki aiheeseen liittyvät taustatiedot sekä kuvaillaan mihin osiin urheiluanalytiikka jakautuu ja käydään läpi myös urheiluanalytiikan tämän hetkisiä haasteita. Luvussa 4 esitetään tutkimuksen tulokset eli kuvaillaan, mihin eksplisiittistä tietoa hyödynnetään koripallossa. Tämän jälkeen luvussa 5 on yhteenveto työstä ja päätelmät, jossa kootaan yhteen päätulokset ja pohditaan niiden merkitystä.

2. TUTKIMUSMENETELMÄ JA -AINEISTO

Tässä luvussa esitellään tutkimusmenetelmä ja -aineisto. Ensimmäisessä alaluvussa käydään tarkemmin läpi, mitä menetelmää tutkimuksessa on käytetty, mistä aineistoa on etsitty ja minkälaista aineistoa on etsitty. Toisessa alaluvussa esitellään tarkemmin tutkimusaineistoa sekä sen keräämiseen liittyviä hakutermejä ja -tuloksia.

2.1 Tutkimusmenetelmä

Tämän kandidaatintyön tutkimusmenetelmänä on kirjallisuuskatsaus. Työssä vertaillaan ja kootaan yhteen aiheesta aikaisemmin tehtyjä tutkimuksia ja muuta tieteellistä kirjallista materiaalia. Tutkimusaineistoa kerätään aluksi yhteen, minkä jälkeen eri lähteistä saatua tietoa hyödynnetään tutkimusongelman ratkaisussa. Työhön etsittiin tieteellistä aineistoa kirjoista, artikkeleista sekä aikaisemmista tutkimuksista.

Aiheesta on saatavilla paljon modernia, 2010-luvun jälkeen julkaistua aineistoa, mikä on hyvä asia tutkimuksen kannalta. Kuten aiemmin mainittiin, urheiluanalytiikka on murrosvaiheessa, joten uusia menetelmiä eksplisiittisen tiedon hyödyntämiseen koripallossa tulee koko ajan lisää. Tämän takia uudet julkaisut ovat hyödyllisiä tutkimusongelman selvittämiseen.

2.2 Tutkimusaineisto

Tutkimusaineistoa etsittiin monista eri tieteellisistä tietokannoista. Tutkimuksessa hyödynnettyjä tietokantoja olivat Andor, Scopus, Web of science ja Google Scholar. Eniten aineistoa tutkimusaiheen kannalta löytyi Scopuksesta sekä Google Scholarista. Suurin osa tutkimuksessa käytetystä aineistosta on artikkeleja, mutta kirjojen tietojakin on hyödynnetty. Tutkimusaineistoksi yritettiin löytää perustietoa urheiluanalytiikasta koripallossa sekä spesifimpää tietoa eksplisiittisen tiedon hyödyntämisestä koripallossa. Tiedonhaussa käytettiin apuna Boolean operaattoreita, jotka toimivat useissa monitieteellisissä tietokannoissa. Aineiston hankinnan kannalta tietokannoista etsittiin materiaalia muun muassa seuraavilla hakutermeillä:

- "Basketball analytics"
- "Basketball analysis"
- Basketball AND "Sport analytics"
- Basketball AND "Sport analysis"
- Basketball AND "Data analytics"

- Basketball AND "Data analysis"

Lähteitä löytyi kyseisillä hakutermeillä kiitettävästi. Osaan mahdollisista työn kannalta hyödyllisistä lähteistä Tampereen Teknillisellä Yliopistolla ei ollut oikeuksia. Aineistoa etsittiin myös tarkennetuimmilla hakusanoilla, mutta yllä olevilla hakutermeillä materiaalia löytyi parhaiten.

3. URHEILUANALYTIikka

Tässä luvussa esitellään tarkemmin urheiluanalytiikan teoreettinen tausta sekä sen nykytila. Urheiluanalytiikkaan, kuten muunlaiseenkin analytiikkaan, kuuluu datan hankkiminen, datan jalostus tiedoksi ja tiedon hyödyntäminen (Shields 2016). Urheiluanalytiikassa hyödynnetään niin hiljaista kuin eksplisiittistäkin tietoa, mutta tässä kandidaatintyössä tutkimuskohteeksi on rajattu vain eksplisiittinen tieto. Tämän luvun tarkoituksena on syventää lukijan tietoisuutta urheiluanalytiikasta ja esitellä kaikki osa-alueet, joita siihen liittyy. Tässä luvussa käydään läpi myös urheiluanalytiikan haasteita koripallon osalta. Tiedon hyödyntämisen keinoja käydään läpi vielä tarkemmin luvussa 4, jossa käydään läpi kirjallisuustutkimuksen avulla saadut ratkaisut tutkimusongelmaan.

Koripallossa on käytetty jo vuosia dataa hyödyksi joukkueen ja yksilöiden suorituksen parantamiseksi. Koripallon pelaamisessa on yksi tarkoitus ja se on vastustajan voittaminen ja tämän saavuttamiseksi on otettu avuksi urheiluanalytiikka. (Gudmundsson & Horton 2017) Aikaisemmin dataa on kerätty lähinnä vain ottelutilastojen muotoon, mutta näistä tilastoista ei saada tarpeeksi hyötyä. Ottelutilastojen on osaltaan huomattu suosivan tietyn tyyppisiä pelaajia, ja heidän todellinen arvo ei tule ilmi näitä tilastoja tutkitessa. (Vaz De Melo et al. 2012) Tämän takia nykyaikana analyttikot joukkueiden taustalla eivät käytä hyödykseen vain ottelutilastoja, vaan muitakin keinoja käytetään hyödyksi (Lucey et al. 2013), josta tarkempaa tietoa tulevissa kappaleissa.

3.1 Datan hankinta ja jalostus

Dataa kerätään useista lähteistä koripallon pelaamisesta. Osa lähteistä on perinteisiä lähteitä, kuten ottelutilastoja, osa lähteistä puolestaan on teknologian tuomien menetelmien kautta saatuja lähteitä, kuten kameroiden keräämää dataa sekä sosiaalisesta mediasta saatavaa dataa. Tässä alaluvussa paneudutaan datan hankintaan ja jalostukseen. Luvussa esitellään laajasti datan lähteet ja kerrotaan lyhyesti myös siitä, miten dataa jalostetaan.

Perinteisin ja vanhin datan keräämiskeino on ottelutilastojen kerääminen. Näitä ottelutilastoja ovat muun muassa pisteiden tekijät, levypallot, syötöt ja heittomäärät. Ottelutilastot ovat siis ihmisten keräämää dataa ottelun tapahtumista. Näitä ottelutilastoja on kerätty pitkään ja niitä on vapaasti saatavilla, mikä helpottaa datan hankintaa. Ottelutilastojen muodossa kerättyä dataa on myös tallessa menneiltä vuosilta. (Vracar et al. 2016) Tämä helpottaa esimerkiksi vuosien varrella tapahtuneen koripallon pelaamisen kehittymisen tutkimista. Ottelutilastot eivät ole kaiken kattavia analysoinnin osalta, minkä takia muutakin dataa tarvitaan. Ottelutilastot eivät yksinään riitä tarpeeksi katta-

vaan analysointiin vaan ne ovat osittain puutteellista tietoa (Vaz De Melo, P O S et al. 2012).

Yksi merkittävä puute ottelutilastoissa on se, että puolustuksesta on liian vähän dataa saatavilla eli suurin osa datasta kerätään hyökkäyspelaamisesta (Franks et al. 2015b). Vaikka ottelutilastoissa on puutteita, niitä tarvitaan urheiluanalytiikassa vieläkin (van Bommel & Bornn 2017). Nykyaikana teknologia mahdollistaa paljon erilaisia asioita, jotka eivät ole ennen olleet mahdollisia. Näitä asioita on myös tullut urheiluanalytiikkaan koripallon osalta. Vuonna 2013 alettiin puhua siitä, että NBA-seurat astuvat big datan aikakaudelle (Goldsberry & Weiss 2013). Tämä johtui siitä, että vuonna 2014 kaikkiin NBA joukkueiden halleihin asennettiin kuusi Sport VU -kameraa, jotka tallentavat lähes kaiken mitä kentällä tapahtuu, myös sen, mitä puolustuspäässä tapahtuu (Sellitto 2015). Nämä kamerat tallentavat pelaajien sekä pallon liikkeitä ja keräävät jopa 25 data pistettä sekunnissa (Sellitto 2015).

Näiden kameroiden datan avulla saadaan tietää, missä ja miten joku asia tapahtui (Lucy et al. 2013). Tämä tieto auttaa täydentämään ottelutilastoja, jotka kertovat ainoastaan, että jotain tapahtui, esimerkiksi piste syntyi ajassa 08:00. Perinteisten tilastojen avulla saadaan tietoa esimerkiksi heitoista, heiton tyypistä, heiton lopputuloksesta ja levypalloista (Bhandari et al. 1997), mutta kameroiden avulla saatu tieto täydentää tätä tietoa tarkemmaksi. Viime vuosikymmenen muutoksena urheiluanalytiikan osalta on myös se, että play-by-play -datan hankinta on kasvanut (Vracar et al. 2016). Tämä tarkoittaa sitä, että yksittäisistä ottelutilanteista on tarkempaa tietoa saatavilla.

Yhtenä teknologian tuomana datan hankinnan lähteenä on myös GPS-laitteiden avulla saatava data. GPS-laitteiden avulla dataa voidaan kerätä pelaajien sijainnista ja liikkeitä kentällä. GPS-laitteita käytetään keräämään dataa harjoituksissa ja peleissä. Tämä on kuitenkin vielä yleisempää harjoittelun puolella kuin otteluiden aikana. Näitä laitteita voidaan hyödyntää esimerkiksi puettavan teknologian avulla eli laittamalla GPS-laitteita pelaajien vaatteisiin. (Therón & Casares 2010).

Datan hankinnasta on huomattava se, että datan lähteet tukevat toisiaan ja täydentävät muista lähteistä saatavan datan aukkoja. Datan hankinnan lisäksi analytiikassa tärkeää on myös sen jalostaminen ja hyödyntäminen. Koripallon osalta urheiluanalytiikkaa käytetään moneen tarkoitukseen ja datan jalostaminen on merkittävä osa sitä. Tiedon hyödyntämistä edeltävässä osassa eli datan jalostuksessa on tärkeää saada data hyödynnettävään muotoon. Urheiluanalytiikassa koripallon osalta datan jalostus on monesti tilastollista mallintamista ja datan visualisointia ymmärrettävään muotoon (Franks et al. 2015a).

Mahdollisena datan lähteenä voidaan pitää myös sosiaalista mediaa. Pelaajat keskustelvat sosiaalisen median, kuten Twitterin, kautta fanien ja muiden toimintaympäristössä toimivien kanssa. Pelaajat julkaisevat erilaisia päivityksiä peleihin tai muuhun elämään-

sä liittyen. Pelaajat käyttävät sosiaalista mediaa mielipiteiden ilmaisuun, mutta päivitysten seurauksia ei mietitä. Pelaajien julkaisuja voidaan analysoida ja hankkia tietoa näiden avulla. Julkaisuissa on aina jokin tunnelataus ja analytytikot voivat hankkia tietoa esimerkiksi pelaajan tunnetilasta, ennen peliä julkaistusta sosiaalisen median päivityksestä. (Xu et al. 2015)

3.2 Tiedon hyödyntäminen

Datan keräämisen ja jalostamisen jälkeen analytiikassa hyödynnetään saatua tietoa. Koripallossa hyödynnetään monesta eri lähteestä saatua eksplisiittistä tietoa monenlaiseen tarkoitukseen. Tässä aluvuossa käydään läpi, mihin tietoa urheiluanalytiikassa koripallon osalta käytetään. Eksplisiittisen tiedon hyödyntäminen on kuitenkin tutkimuksen pääkohteena ja näin ollen tätä osaa käsitellään myöhemmässä vaiheessa tarkemmin. Urheiluanalytiikkaan liittyy lukuisia haasteita, joiden ratkaisuun hankitaan eksplisiittistä tietoa ja pyritään myös hyödyntämään sitä.

Eksplisiittistä tietoa hyödynnetään päätöksentekoprosessissa. Koripallon pelaamisen osalta päätöksiä joudutaan tekemään koko ajan ja urheiluanalytiikan tarkoituksena on olla tukemassa päätöksentekoa. (Shields 2016) Urheilusuorituksen kannalta päätöksiä tekee valmentajat ja pelaajat. Sekä valmentajat, että pelaajat joutuvat koko ajan pohtimaan, mistä ratkaisusta joukkue hyötyisi eniten. Urheiluanalytiikan ja erityisesti eksplisiittisen tiedon avulla voidaan arvioida, mikä päätöksistä toisi eniten arvoa kussakin tilanteessa (Cervone et al. 2014).

Tietoa hyödynnetään yleisesti myös pelistrategian suunnittelussa (Goldsberry & Weiss 2013). Strategiaa suunnitellaan esimerkiksi sen mukaan, millainen vastustaja on ja ketä omassa joukkueessa pelaa. Vastustajan analysointiin liittyy vahvasti myös tulevan ennustaminen, johon urheiluanalytiikan avulla eksplisiittistä tietoa voidaan hyödyntää (Yue et al. 2015). Tällä tulevan ennustamisella tarkoitetaan esimerkiksi sitä, että tiedetään, mitä vastustaja yleensä tekee ja miten heidät voisi voittaa. Vaikka nykyisin apuna on paljon dataa ja siitä jalostettua eksplisiittistä tietoa, tulevan ennustaminen on silti hankalaa koripallon pelaamisessa (Gudmundsson & Horton 2017).

Koripallon pelaamisen päällimmäisenä tavoitteena on voittaa vastustaja ja voittoa yritetään tavoitella urheiluanalytiikan keinoin. Kuten aiemmin mainittiin, urheiluanalytiikkaa hyödynnetään strategisessa suunnittelussa. Näin ollen urheiluanalytiikkaa hyödynnetään parantamaan voiton todennäköisyyttä (Vracar et al. 2016). Voiton todennäköisyyttä yritetään parantaa strategian suunnittelun lisäksi valitsemalla oikeat pelaajat joukkueeseen ja laittamalla heidät oikeisiin rooleihin. Tietoa voidaan hyödyntää pelaajien vertailuun ja heidän todellisen arvon määrittelyyn (Vinué & Epifanio 2017). Urheiluanalytiikan avulla pelaajia voidaan vertailla esimerkiksi heidän fyysisten ominaisuuksiensa osalta (Balli & Korukoglu 2014). Merkittäviä eksplisiittisen tiedon hyödyntämisen kohde on myös loukkaantumisten analysointi.

Urheiluanalytiikkaa käytetään loukkaantumisten tutkimiseen. Loukkaantumisista on kerätty pitkään dataa ja loukkaantumistilastoja analysoidaan. Tilastojen pohjalta eri lajien, kuten koripallon, eniten esiintyvät loukkaantumiset ovat selvitettävissä. Loukkaantumisten esiintyessä, tämän tiedon johdosta eri urheilulajien sääntöjä on tiukennettu. Sääntöjä on tiukennettu, koska halutaan välttää turhilta loukkaantumisilta ja tämä voidaan estää tiukemmilla säännöillä. (Kujala et al. 1995). Esimerkiksi koripallon osalta eniten loukkaantumisia on polvissa ja nilkoissa. Tämä tieto mielessä pitäen voidaan yrittää varoa tilanteita, joissa loukkaantumisia voi sattua.

3.3 Urheiluanalytiikan haasteet

Urheiluanalytiikkaan koripallon osalta liittyy ja on liittynyt useita haasteita vuosien varrella. Tässä luvussa esitellään kaikki koripalloon liittyvät haasteet, joihin modernilla urheiluanalytiikalla voidaan vastata ja, joita urheiluanalytiikka on tuonut mukanaan. Haasteina urheiluanalytiikassa koripallon osalta ovat puolustuksen vähäinen analysointi, vanhoihin pelaajien rooleihin tukeutuminen, pelaajien todellisen arvon määrittäminen, tuurin olemassaolo kaikissa koripallo-otteluissa sekä sosiaalisen median ja uusien teknologioiden tuomat haasteet.

Koripalloa on pelattu pitkään ja siitä on kerätty dataa jo vuosia. Ensimmäiset datat ovat kerätty jo 1860-luvulla. (Gudmundsson & Horton 2017) Dataa on siis kerätty pitkään, mutta sitä ei ole hyödynnetty datan määrään nähden tarpeeksi. Koripallon osalta tietoa ei ole hyödynnetty laajasti ennen kuin on huomattu sen tuovan tarpeeksi etua. Ensimmäisiä merkittäviä urheiluanalytiikan tunnettuja hyödyntäjiä olivat NBA:ssa Dallas Mavericks sekä Golden State Warriors. Vuonna 2011 Dallas Mavericks voitti NBA-mestaruuden, ja heillä oli käytössään sinä vuonna urheiluanalytiikan avulla kehitetty strategia, jossa heidän taktiikkanaan oli liikuttaa paljon palloa ja saada se loppujen lopuksi parhaalle heittäjälle (Fewell et al. 2012). He olivat tutkineet joukkuettaan verkostona ja tämän avulla luoneet uuden strategian (Fewell et al. 2012).

Toisena urheiluanalytiikan kannalta tärkeänä edistäjä toimi Golden State Warriors, joka on voittanut viime vuosina monta kertaa NBA-mestaruuden. Heitä pidetään yhtenä urheiluanalytiikan hyödyntämisen pioneereinä, ja he ovat hyödyntäneet dataa päätöksenteossa (Shields 2016). Osittain näiden joukkueiden menestyksen ansiosta joukkueet hyödyntävät entistä enemmän olemassa olevaa dataa. Urheiluanalytiikan osalta perinteisimpiä keinoja ja datan lähteitä on ollut ottelutilastojen kerääminen sekä hyödyntäminen. Ottelutilastoja on kerätty pääosin vain hyökkäämispelaamisesta, jolloin puolustuspeleamisen statistiikka on jäänyt vähälle ja vain helposti huomattavia asioita on arvioitu puolustuksesta (Franks et al. 2015).

Urheiluanalytiikka ei ole siis puolustuksen osalta tarpeeksi hyvällä tasolla (Goldsberry & Weiss 2013). Koripallon perusideana on tehdä koreja hyökkäyspäädyssä ja puolustaa vastustajaa tekemästä koreja omassa puolustuspäädyssä. Hyökkäyksen analysointiin ja

tilastointiin on panostettu NBA:ssa, mutta puolustuksesta on tilastoitu vain perustilastoja, kuten plus/minus-tilastoja ja levypalloja. Tämän kaltaisten tilastojen avulla on vaikea selvittää, miksi joku pelaaja on tehokas puolustaja tai onko hän oikeasti edes tehokas vai onko hän ainoastaan kerättyjen tilastojen mukaan tehokas. (Franks et al. 2015) Tämä on koripallojoukkueiden kannalta suuri haaste, johon urheiluanalytiikalla voidaan vastata.

Puolustuspelaamisen arvioinnin vaikeus ei ole kuitenkaan ainut haasteellinen arvioinnin kohde koripallopelaajissa. Pelaajan todellisen arvon selvittäminen on tärkeää joukkueen kannalta (Vinué & Epifanio 2017). Tämä tarkoittaa pelaajan arviointia kokonaisuudessa eli sitä, kuinka arvokas hän todellisuudessa on. Perinteisten tilastojen pohjalta ei ole mahdollista arvioida pelaajaa tarpeeksi tarkasti, eikä näin ollen osata sanoa, onko pelaaja palkkansa arvoinen (Fewell et al. 2012). Pelaajien arviointia ja heidän vertailua vaikeuttaa myös se, että perinteiset viisi roolia ovat modernissa koripallon pelaamisessa muuttuneet todellisuudessa, jopa 13 rooliin (Bianchi et al. 2017).

Rooleja on aina ollut koripallossa, ja erilaiset pelaajatyypit ovat pelanneet eri rooleissa. Rooleihin jakamisen avulla pelaajien vertailu ja strategian luonti on ollut helpompaa. Koripallon pelaamisen muututtua pelaajatyypit ovat myös muuttuneet, ja näin ollen muutos on luonut uusia rooleja. Koripallon pelaaminen keskittyy nykyään enemmän kolmen pisteen heittoihin, kuin ennen, koska on huomattu niiden tuovan paljon arvoa (Goldsberry & Weiss 2013). Uudenlaisen pelitavan takia pelaajien vertailu keskenään vanhojen roolien valossa on haaste, sillä roolit eivät todellisuudessa ole vertailukelpoisia enää. (Bianchi et al. 2017) Koripallossa pelaamisen muutoksen lisäksi sen toimintaympäristö on muuttunut. Kaikkien joukkueiden halleihin asennettujen kameroiden vuoksi dataa on saatavilla entistä enemmän, ja tämä on ratkaissut joitain haasteita, mutta tuonut mukanaan myös uusia haasteita.

Aikaisemmin, kun dataa kerääviä kameroita ei ollut, ihmiset joutuivat mekaanisesti ottamaan ylös ottelun tapahtumat. Tämän haasteena on se, että ihminen onnistuu keräämään ainoastaan 42 % pelin tapahtumista (Gudmundsson & Horton 2017). Toisena haasteena mekaanisessa tiedon keräämisessä on ollut myös se, että ottelutilastot kertovat ainoastaan, mitä on tapahtunut. Tällainen tieto ei kuitenkaan riitä, vaan nykyään halutaan tietää, missä on tapahtunut ja miten tapahtumaan on päädytty (Lucey et al. 2013). Esimerkkinä tästä joukkueet eivät halua ainoastaan tietää, että he tekivät kahden pisteen heiton kolmannen neljänneksen alussa, vaan halutaan tietää, mistä heitto lähti ja miten tilanteeseen päädyttiin. Kameroiden takia myös valmentajien kykyjä haastetaan uudella tavalla. Valmentajilta vaaditaan nykyään visuaalista muistia, jotta he muistavat miten tietyt liikkeet ja kuviot toimivat tietyissä tilanteissa (Therón & Casares 2010).

Urheiluanalytiikan osalta on kuitenkin tärkeä muistaa, että sattuma on aina läsnä otteissa ja tämä tuo haasteita tulosten ennustamiseen. Korkealla tasolla pelaavat, kuten NBA:ssa pelaavat, pelaajat harjoittelevat tiettyjä asioita ja suorittavat tiettyjä liikkeitä

ottelusta toiseen. Pelaajat siis pelaavat tietyllä odotetulla tavalla, mutta tuurin ollessa mukana lopputuloksen analysointi on vaikeaa. (Clauzet et al. 2015) Tämä ennustamisen haaste näkyy muun muassa siinä, että keskimääräisesti katsottuna altavastaajajoukkueen todennäköisyys voittaa NBA-ottelu on 0,36, mutta tuurin astuessa kuvaan altavastaajajoukkueetkin voittavat otteluita (Aoki et al. 2017).

4. EKSPLISIITTISEN TIEDON HYÖDYNTÄMINEN KORIPALLOSSA

Tässä luvussa käydään läpi eksplisiittisen tiedon hyödyntämisen kohteita koripallossa. Aihetta pohjustettiin luvussa 3.2, mutta tämän luvun tarkoituksena on kuvailla tarkkaan kaikki eksplisiittisen tiedon käyttökohteet. Tässä luvussa käydään läpi tulokset ja ratkaisut, joita on saatu tutkimusongelmaan sekä tutkimuskysymyksiin. Luvussa kerrotaan siis, miten eksplisiittistä tietoa voidaan hyödyntää urheilusuorituksen parantamiseen koripallon pelaamisessa. Alaluvut ovat jaettu pienempiin kokonaisuuksiin eri tiedon lähteiden tai käyttökohteen mukaan.

4.1 Ottelutilastot

Ottelutilastojen hyödyntäminen koripallossa on urheiluanalytiikan vanhin muoto. Kuten työssä on aiemminkin mainittu, ottelutilastoja on kerätty 1860-luvulta saakka (Gudmundsson & Horton 2017). Ottelutilastoja on laajasti saatavilla ja niitä voidaan käyttää moneen tarkoitukseen. Statistiikasta ja matemaattisesta mallinnuksesta on tullut tärkeä osa urheilua ja sen eteen nähdään paljon vaivaa (Vracar et al. 2016). Ottelutilastot eivät ole kaikista parhaimpia urheilusuorituksen mittareita, mutta silti niitä käytetään vielä moniin tarkoituksiin (Vaz De Melo et al. 2012).

Urheiluanalytiikan alkutaipaleella ottelutilastot olivat ainoa asia, jota hyödynnettiin ja kerättiin (Lucey et al. 2013). Nykyään ottelutilastoja hyödynnetään pitkälti yhdistettynä muuhun dataan. Ottelutilastojen yhtenä ongelmana koripallon osalta on se, että näitä tilastoja kerätään pääosin vain hyökkäyspelaamisesta (Franks et al. 2015). Tämä on ongelma muun muassa sen takia, että pelaajia arvioidaan ja heidän arvoa mitataan pääosin vain hyökkäystaitojen perusteella (Franks et al. 2015). Ottelutilastojen yksi keskeinen käyttökohde urheiluanalytiikassa koripallon osalta on juuri pelaajien arviointi ja vertailu.

Ottelutilastojen avulla voidaan selvittää, mitä kentällä on tapahtunut, mutta ei sitä, miksi jotain on tapahtunut (Franks et al. 2015). Ottelutilastoja käytetään siis tilastoimaan kentällä tapahtuneita asioita ja näiden tilastojen avulla on mahdollista esimerkiksi vertailla pelaajia keskenään ja arvioida heidän suorituksiaan (van Bommel & Bornn 2017). Yhdistämällä ottelutilastoja kameroiden avulla kerättyyn dataan saadaan vielä tarkempaa ja syvällisempää tietoa kentän tapahtumista (Franks et al. 2015). Näiden yhdistelmä tietojen avulla esimerkiksi puolustuspelaamista on tarkempi arvioida, ja pelaajien todellinen arvo on paremmin määriteltävissä.

Ottelutilastoja on helposti saatavilla ja tämä on yksi syy sen suuren käyttöön. Esimerkiksi NBA-Stats –internetsivustolla (NBA Stats) on kerätty tilastoja perinteisistä asioista ja tämän lisäksi sieltä löytyy myös Advanced stats –osio, jossa on kehittyneempiä tilastoja. Perinteisimpiä tilastoja ovat pisteet, levypallot, syötöt, blokkaukset, riistot, vastahyökkäykset ja heittoprosentit (NBA Stats). Ottelutilastoja on siis laajalta alueelta saatavilla ja tämä kertoo siitä, että jo pelkästään niiden avulla on mahdollista analysoida pelaajia hyvinkin pitkälle.

Kuten aiemmin mainittiin, ottelutilastoja käytetään pelaajien arviointiin. Ottelutilastot ovat tärkeässä roolissa, kun määritellään pelaajien pelaikoja, palkkoja ja heidän markkina-arvoa (van Bommel & Bornn 2017). Käytännössä ne pelaavat, jotka ovat tilastojen valossa arvokkaita joukkueelle voittamisen kannalta, ja tilastot määrittelevät myös osittain heidän palkkansa sekä sen, minkä arvoisia he ovat pelaajamarkkinoilla. Vaikka ottelutilastoja käytetään pelaajien arvon määrittelyyn, on tärkeä huomata, että ne eivät kuitenkaan kerro pelaajan todellista arvoa, eikä sitä, kuinka arvokas pelaaja on joukkueelle voittamisen kannalta (Deshpande & Jensen 2016).

4.2 Seurantadata

Työssä jo moneen kertaan mainitut Sport VU -kamerat tuottavat paljon dataa koripallon pelaamisesta. Kuusi kameraa jokaisen NBA-joukkueen hallissa tallentavat jopa 25 datamittausta sekunnissa (Sellitto 2015). Tämä tarkoittaa todella suurta datamäärää, kun huomioidaan, että NBA-sarjassa on 30 joukkuetta ja jokainen joukkue pelaa runkosarjassa 82 ottelua, monilla joukkueilla tähän pelimäärään lisätään vielä pudotuspelit (NBA.com). Nämä kamerat tuottavat siis seurantadataa pelaajien ja pallon liikkeistä eli kentän tapahtumista (Gudmundsson & Horton 2017).

Tätä dataa jalostetaan haluttuun muotoon ja jalostettua eksplisiittistä tietoa hyödynnetään moneen tarkoitukseen. Urheiluanalytiikassa koripalloilussa halutaan tietää, missä ja miten jokin asia tapahtui kentällä (Lucey et al. 2013). Tähän kysyntään voidaan vastata ottamalla ottelutilastojen tueksi seurantadataa. Pelaajia sekä palloa seurataan kameroiden avulla koko ajan, ja tämän ansioista on helpompaa selvittää syitä tapahtuneille asioille. Kuten aiemmin on mainittu, tulevan ennustaminen on haasteellista koripallossa, etenkin pelkkien ottelutilastojen avulla. Tähän tulevan ennustamiseen voidaan hyödyntää seurantadataa (Yue et al. 2015).

Pelaajien sekä joukkueiden hyvään analysointiin tarvitaan tietoa pelaajien liikkeistä ja kanssakäymisistä toisten pelaajien kanssa. Pelaajien välisen kanssakäymisten ja toimivien yhdistelmien analysointi on vaikein osa-alue seurantadataa hyödyntävistä kohteista. Seurantadatan avulla on analysoitu eniten pelaajien syöttelyä ja liikkumista. Seurantadatan analysoinnin avulla voidaan esimerkiksi arvioida heiton riskiä tai syötön laatua, mutta tulevan ennustaminen on vielä haastavaa tämän datan pohjalta. Seurantadatan analysoinnin avulla luodaan tietoa siitä, milloin kentällä kannattaa tehdä mitäkin ja ke-

nen kannattaa tehdä mitään. (Gudmundsson & Horton 2017) Esimerkiksi kannattaako kolmen pisteen heittoa heittää, jos omia pelaajia ei ole korin läheisyydessä vai kannattaisiko siinä tilanteessa mieluummin syöttää tai ajaa korille.

Seurantadatan avulla pelaajien tehokkuuksia voidaan myös arvioida. Hyökkäys- tai puolustuspään tehokkuutta voidaan analysoida tutkimalla sitä, ketkä ovat olleet kentällä, kun hyökkäys tai puolustus on ollut tehokasta, ja mitä silloin on tehty hyvin tai huonosti. (Gudmundsson & Horton 2017) Ottelutilastojen tutkiminen ei ole paras pelaajien arvioinnin keino, sillä se mittaa vain tiettyjä spesifejä asioita, esimerkiksi pisteiden määrää. Pisteiden tekeminen ei ole kuitenkaan strategian kannalta aina järkevintä vaan muut ratkaisut, kuten syöttäminen, voi tuoda enemmän arvoa joukkueelle (Cervone et al. 2014).

Perusmittareiden tueksi on luotu EPV-mittari, jonka avulla voidaan tutkia pelaajien todellista arvoa esimerkiksi hyökkäyspelaamisessa. EPV-termi tulee englanninkielen sanoista *expected possession value*, joka tarkoittaa käytännössä pelaajan luomaa arvoa pallon ollessa hänellä. EPV-mittari muodostetaan hyödyntäen seurantadataa. (Cervone et al. 2014). Tämän mittarin avulla pelaajien arviointi ja keskinäinen vertailu ovat helpompaa ja valmentaja tietävät, ketkä pelaajista ovat arvokkaita joukkueelle. Urheiluanalytiikan tarkoituksena on tukea päätöksen tekoa, ja näiden tietojen pohjalta valmentajan on helpompi tehdä ratkaisuja.

Pelistrategiaan liittyen seurantadatan avulla on saatu tietää, että NBA:ssa keskimääräinen heittoprosentti on 40 %, jos pääsee vapaasti heittämään ja 32 %, jos ei pääse vapaasti heittämään. Kameroiden datan jalostuksen avulla on mahdollista tutkia, miten heittopaikkoihin on päästy. Esimerkiksi pelintekijän ja pienen laitahyökkääjän paikkojen vaihto on monesti luonut tilanteen, jossa päästään heittämään vapaasti. (Lucey et al. 2014) Tämä tieto on merkittävää, sillä strategiaa voidaan luoda sen pohjalta, miten päästään parhaiten vapaaseen heittopaikkaan. Kahdeksan prosentin ero on todella suuri, sillä koripallossa tulee paljon heittoa ottelun aikana, jolloin etu kasvaa suureksi.

4.3 Roolien muuttuminen

Koripallossa on aina ollut eri rooleja ja pelipaikkoja. Perinteisesti katsottuna näitä rooleja on viisi. Nämä viisi pelipaikkaa ovat pelintekijä, heittävä takamies, pieni laitahyökkääjä, iso laitahyökkääjä ja sentteri (Koripallo.net). Pelipaikoilla on kullakin omat tehtävänsä ja tietynlaiset pelaajat ovat sopineet tietynlaiseen pelipaikkaan. Datan avulla nykyään tiedetään, että tietyn tyyppiset pelaajat pelaavat paremmin yhteen tietyn tyyppisten pelaajien kanssa, ja tätä voidaan hyödyntää pelitaktiikkaa muodostettaessa (Lutz 2012).

Kuten edellisessä kappaleessa kerrottiin, rooleja on aina ollut. Roolit tai pelaajien jaottelu pelipaikkoihin on helpottanut pelaajien hankintaa ja vertailua. Nykyisen tietämyksen

ja datan pohjalta on kuitenkin huomattu, että nämä perinteiset pelipaikkajaottelut eivät ole tarpeeksi kattavalla tasolla moderniin koripalloon nähden. Topologisen analytiikan avulla on selvitetty, että pelaajatyyppejä on oikeasti 13, eikä vain viisi. Tämän uuden pelipaikkajaottelusta saadun tiedon avulla voidaan entistä tarkemmin suunnitella joukkueen taktiikkaa, esimerkiksi hyökkäyskuviota. Uusien roolien myötä analysoinnin avulla on myös selvitetty, mitkä roolit mihinkin strategiaan sopivat ja, minkä tyyppisten pelaajien kannattaa tehdä mitäkin kuvioita. Uuden roolijaottelun avulla myös pelaajien vertailu on entistä tarkempaa. (Bianchi et al. 2017)

4.4 Teknologian tuomat mahdollisuudet

Nykytekniikka mahdollistaa monenlaisia mahdollisuuksia datan hankinnalle ja tiedon hyödyntämiselle. Yksi koripallossa käytetty urheiluanalytiikan keino on GPS-laitteiden hyödyntäminen. GPS (Global positioning system)-laitteiden ansiosta dataa voidaan kerätä pelaajien sijainnista kentällä (Therón & Casares 2010). GPS-laitteita hyödynnetään niin harjoittelussa kuin peleissäkin, mutta se on yleisempää harjoittelussa. GPS-laitteita, jotka keräävät ja luovat dataa ovat muun muassa puettava teknologia eli urheilijan vaatteisiin asennettu laite, jossa on GPS-sensori. Näiden laitteiden seurattessa pelaajia on esimerkiksi valmentajien helpompi vertailla pelaajia. (Therón & Casares 2010) Esimerkiksi pelaajien työmäärää voidaan vertailla ja tästä saadaan tietää, kuka pelaaja tekee minkäkin verran työtä kentällä, tai pelaajan omia työmääriä voidaan vertailla ja tehdä tästä analyysseja.

Työmääristä voidaan selvittää myös se, onko pelaaja palautunut hyvin otteluista ja harjoituksista vai onko hän väsynyt. Työmäärän ollessa matalalla tasolla pelaajan omaan tasoon nähden voi tilanne johtua väsymyksestä, joka taas puolestaan johtaa loukkaantumisiin. Loukkaantumisten analysointia on tehty jo pitkään urheilun osalta (Kujala et al. 1995) ja teknologia on valmiina GPS-laitteiden laajempaan hyödyntämiseen (Therón & Casares 2010). Näiden GPS-laitteiden avulla voitaisiin esimerkiksi kehittää mittareita, jonka avulla loukkaantumisten määrä vähenisi.

4.5 Sosiaalinen media

Sosiaalinen media on tuonut paljon uusia asioita elämäämme viime vuosina. Se on tuonut myös koripalloon mahdollisuuden saada tietoa pelaajista. Monet koripalloilijat käyttävät sosiaalista mediaa, kuten Twitteriä ja julkaisevat sinne ajatuksiaan. Pelaajat käyttävät esimerkiksi Twitteriä keskusteluun fanien, muiden pelaajien ja organisaatioiden kanssa eikä sanomisien seurauksia mietitä tarkkaan (Xu et al. 2015). Kaikissa sosiaalisen median julkaisuissa on aina jonkinlainen tunnelataus mukana, positiivinen tai negatiivinen.

Hyödyntäen pelaajien julkaisuja, voidaan analysoida heidän tunnetilaansa ja ennustaa heidän suoriutumistaan kentällä. Valmentajat, joukkueiden toimitusjohtajat sekä muu

henkilökunta voivat hyödyntää sosiaalisen median julkaisujen analysointia havaitakseen ja tunnistaakseen pelaajan tunnetiloja ennen peliä. Valmentajat, pelaajat, fanit ja myös tutkijat uskovat siihen, että kentän ulkopuoliset tunnetilat vaikuttavat pelaajien suorituksiin pelikentillä. (Xu et al. 2015)

5. YHTEENVETO

Tässä luvussa käydään läpi tutkimuksen yhteenveto, sekä tulosten arviointi ja jatkotutkimusmahdollisuudet. Yhteenvedossa kootaan päätulokset yhteen tiiviisti. Tulosten arvioinnissa ja jatkotutkimus mahdollisuuksissa arvioidaan saatuja tuloksia ja pohditaan mihin suuntaan tutkimusta voisi jatkaa.

5.1 Yhteenveto

Tässä kandidaatintyössä tutkittiin eksplisiittisen tiedon hyödyntämistä urheiluanalytiikassa koripallovalmennuksen osalta ja työssä keskityttiin pääosin NBA-koripalloiluun. Tutkimukseen sisältyi urheilu suorituksen parantamisen lisäksi myös muita valmennuksen tehtäviä, kuten pelaajien vertailua. Työ toteutettiin kirjallisuus katsauksena, etsien kirjallista aineistoa useista eri tietokannoista. Tutkimusongelmana työssä oli se, että eksplisiittistä tietoa ei olla hyödynnetty sillä tasolla, kuinka paljon dataa koripallon pelaamisesta oikeasti kerätään. Tietoa ei siis olla hyödynnetty niin paljon, kuin olisi mahdollista. Tutkimuksen pääkysymyksenä työssä oli se, miten eksplisiittistä tietoa voidaan hyödyntää urheilu suorituksen parantamisessa koripallossa.

Työn tulokset osoittavat, että dataa kerätään monista lähteistä ja jalostettua eksplisiittistä tietoa hyödynnetään moniin tarkoituksiin. Dataa kerätään ottelutilastojen avulla, Sport VU –kameroiden ja GPS-laitteiden kautta, sekä sosiaalisen median julkaisuja tutkimalla. Perinteisin datan keräys muoto on ottelutilastot, joita on kerätty jo ennen 1900-lukua. Ottelutilastojen lisäksi merkittävimmän määrän dataa tuottaa Sport VU –kamerajärjestelmät, jotka ovat asennettu kaikkien NBA-joukkueiden halleihin. Nämä kamerajärjestelmät keräävät 25 data pistettä sekunnissa pelaajien ja pallon liikkeistä kentällä. GPS-laitteiden ansiosta on mahdollista saada tietoa pelaajien sijainneista ja liikkeistä. Sosiaalisen median julkaisujen avulla saadaan tietoa pelaajien tunnetiloista, heidän julkaisemien päivitysten kautta.

Datasta jalostettua eksplisiittistä tietoa voidaan hyödyntää moneen tarkoitukseen urheiluanalytiikassa koripalloilun osalta. Nämä hyödyntämisen keinot jaettiin työssä ottelutilastoihin, seurantadataan, roolien muuttumiseen, GPS-teknologiaan ja sosiaaliseen mediaan. Ottelutilastoja hyödynnetään pelitapahtumien läpikäymiseen eli niiden avulla voidaan tutkia, mitä pelissä tapahtui. Ottelutilastojen avulla voidaan arvioida pelaajien suorituksia ja vertailla niitä keskenään. Pelaajien arvo määräytyy myös osittain heidän ottelutilastojen analysoinnin pohjalta. Ottelutilastoissa on kuitenkin puutteita, mutta nykyajan teknikoiden avulla puutteita voidaan täydentää esimerkiksi seurantadatasta jalostetulla tiedolla.

Seurantadasta jalostetun tiedon avulla voidaan tukea ottelutilastojen avulla tehtyä analytiikkaa ja luoda uusia menetelmiä urheiluanalytiikalle. Seurantadatan avulla voidaan selvittää syitä sille, miksi, missä ja miten jokin asia tapahtui kentällä. Tämän avulla voidaan tehdä johtopäätöksiä siitä, mitkä asiat ovat hyvin ja mitkä huonosti. Seurantadattaa hyödynnetään myös tulevan onnistamiseen, joka on kuitenkin haasteellista, sillä koripallon pelaamisessa on aina tuuria mukana. Pelaajien ja joukkueiden tarkempi analysointi on mahdollista, kun ottelutilastojen tueksi tuodaan kamerajärjestelmän avulla saatua seurantadattaa. Tämän seurantadasta saadun tiedon pohjalta joukkueet muodosta strategiansa, sillä saatu tieto kertoo sen, mitkä kuviot ovat toimineet aikaisemmin ja mitkä ovat vastustajan heikkoudet. Seurantadatan avulla voidaan siis arvioida, mitä kannattaa tehdä missäkin tilanteessa, esimerkiksi voidaan arvioida heittojen riskiä ja sitä, missä tilanteessa kannattaa heittää.

Urheiluanalytiikassa koripallon osalta eksplisiittistä tietoa hyödynnetään edellä mainittujen tapojen lisäksi myös muutenkin. Esimerkiksi pelaajien arviointiin hyödynnetään sosiaalisesta mediasta saatua tietoa pelaajien tunnetiloista ja GPS-laitteiden avulla voidaan analysoida pelaajien työmääriä sekä vertailla pelaajien statistiikkaa keskenään. Seurantadatan keräämisen ja urheiluanalytiikan kehittymisen myötä on huomattu että, perinteiset koripalloilun pelaajatyypit eivät modernissa koripallossa pidä täysin paikkaansa. Modernissa koripallossa on urheiluanalytiikan perusteella nykyaikana 13 pelaajatyyppeä, perinteisesti on ajateltu, että pelaajatyyppejä on vain viisi. Uusien pelaajatyyppeiden avulla pelaajien vertailu on tarkempaa ja pelikuvioiden luonti on parempaa, kun tiedetään tarkemmin, minkä tyyppiset pelaajat pelaavat hyvin yhteen.

5.2 Tulosten arviointi ja jatkotutkimusmahdollisuudet

Tutkimuksen tuloksia voidaan pitää oikeellisina, sillä kirjallisia lähteitä vertailtiin keskenään ja niiden väliltä löytyi paljon yhtäläisyyksiä. Tulosten arvioinnin kannalta on tärkeä huomata myös, että suurin osa lähteistä on 2010-luvun aineistoja, joka on tärkeää, sillä suurin urheiluanalytiikan kehitys on vasta tapahtunut tai on vasta tapahtumassa. Tutkimuksessa etsittiin keinoja hyödyntää eksplisiittistä tietoa koripallossa ja tähän kysymykseen löydettiin monta keinoja. Näin ollen voidaan sanoa, että tuloksia syntyi kattava määrä.

Jatkotutkimusmahdollisuutena olisi tärkeää, että tutkimukseen otettaisiin huomioon esimerkiksi valmentajilla olevaa hiljaista tietoa. Tietoa hyödynnetään varmasti todella monella muullakin eri tavoilla, mitä työssä ei tullut vastaan. Tämä tieto saattaa kuitenkin olla tarkoin varjeltua hiljaista tietoa, sillä vastustajille ei haluta kertoa kaikkia omia urheiluanalytiikan keinoja. Jatkotutkimuksien osalta datan hankintaa ja jalostusta voitaisiin tutkia vielä enemmän, sillä tässä työssä kyseiset aihealueet jäivät niukalle huomiolle.

LÄHTEET

Aoki, R., Assuncao, R.M. & Vaz de Melo, Pedro OS (2017). Luck is hard to beat: The difficulty of sports prediction, *ACM*, pp. 1367-1376.

Balli, S. & Korukoglu, S. (2014). Development of a fuzzy decision support framework for complex multi-attribute decision problems: A case study for the selection of skilful basketball players, *Expert Systems*, Vol. 31(1), pp. 56-69.

Bhandari, I., Colet, E., Parker, J., Pines, Z., Pratap, R. & Ramanujam, K. (1997). Advanced scout: Data mining and knowledge discovery in NBA data, *Data Mining and Knowledge Discovery*, Vol. 1(1), pp. 121-125.

Bianchi, F., Facchinetti, T. & Zuccolotto, P. (2017). Role revolution: towards a new meaning of positions in basketball, *Electronic Journal of Applied Statistical Analysis*, Vol. 10(3), pp. 712-734.

van Bommel, M. & Bornn, L. (2017). Adjusting for scorekeeper bias in NBA box scores, *Data Mining and Knowledge Discovery*, Vol. 31(6), pp. 1622-1642.

Cervone, D., D'Amour, A., Bornn, L. & Goldsberry, K. (2014). POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data, pp. 3.

Clauset, A., Kogan, M. & Redner, S. (2015). Safe leads and lead changes in competitive team sports, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, Vol. 91(6),

Deshpande, S.K. & Jensen, S.T. (2016). Estimating an NBA player's impact on his team's chances of winning, *Journal of Quantitative Analysis in Sports*, Vol. 12(2), pp. 51-72.

Fewell, J.H., Armbruster, D., Ingraham, J., Petersen, A. & Waters, J.S. (2012). Basketball teams as strategic networks. *PloS one*, Vol. 7(11).

Franks, A., Miller, A., Bornn, L. & Goldsberry, K. (2015a). Counterpoints: Advanced defensive metrics for nba basketball.

Franks, A., Miller, A., Bornn, L. & Goldsberry, K. (2015b). Characterizing the spatial structure of defensive skill in professional basketball, *Annals of Applied Statistics*, Vol. 9(1), pp. 94-121.

Goldsberry, K. & Weiss, E. (2013). The Dwight effect: A new ensemble of interior defense analytics for the NBA, pp. 1-11.

Gudmundsson, J. & Horton, M. (2017). Spatio-temporal analysis of team sports, *ACM Computing Surveys*, Vol. 50(2).

Koripallo.net. Saatavilla (<http://www.koripallo.net/koripallo/taktiikat/pelipaikat/>). Luettu (27.3.2018).

Kujala, U.M., Taimela, S., Antti-Poika, I., Orava, S., Tuominen, R. & Myllynen, P. (1995). Acute injuries in soccer, ice hockey, volleyball, basketball, judo, and karate: Analysis of national registry data, *BMJ*, Vol. 311(7018), pp. 1465.

Lucey, P., Bialkowski, A., Carr, P., Yue, Y. & Matthews, I. (2014). How to get an open shot: Analyzing team movement in basketball using tracking data.

Lucey, P., Oliver, D., Carr, P., Roth, J. & Matthews, I. (2013). Assessing team strategy using spatiotemporal data, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1366-1374.

Lutz, D. (2012). A cluster analysis of NBA players.

Martinez, J.A., Ruiz, M., Casals, M. & Lopez, F. (2017). Regular point scoring by professional basketball players, *Electronic Journal of Applied Statistical Analysis*, Vol. 10(3), pp. 759-772.

NBA Stats. Saatavilla (<https://stats.nba.com>). Luettu (27.3.2018).

NBA.com. Saatavilla (<http://www.nba.com/#/>). Luettu (27.3.2018).

Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis*, Potomac Books, Inc.

Sellitto, C. (2015). Big Data can give Athletes the Winning Edge, *Scientific Computing*.

Shields, B. (2016). Stephen Curry, the Golden State Warriors, and the Power of Analytics at Work.

Therón, R. & Casares, L. (2010). Visual analysis of time-motion in basketball games, *Springer*, pp. 196-207.

Vaz De Melo, P O S, Almeida, V.A.F., Loureiro, A.A.F. & Faloutsos, C. (2012). Forecasting in the NBA and other team sports: Network effects in action, *ACM Transactions on Knowledge Discovery from Data*, Vol. 6(3).

Vinué, G. & Epifanio, I. (2017). Archetypoid analysis for sports analytics, *Data Mining and Knowledge Discovery*, Vol. 31(6), pp. 1643-1677.

Vracar, P., Štrumbelj, E. & Kononenko, I. (2016). Modeling basketball play-by-play data, *Expert Systems with Applications*, Vol. 44 pp. 58-66.

Xu, C., Yu, Y. & Hoi, C.-. (2015). Hidden in-game intelligence in NBA players' tweets, *Communications of the ACM*, Vol. 58(11), pp. 80-89.

Yue, Y., Lucey, P., Carr, P., Bialkowski, A. & Matthews, I. (2015). Learning Fine-Grained Spatial Models for Dynamic Sports Play Prediction, Proceedings - IEEE International Conference on Data Mining, ICDM, pp. 670-679.